

# Patent Abstracts of Japan

PUBLICATION NUMBER : 10187183  
PUBLICATION DATE : 14-07-98

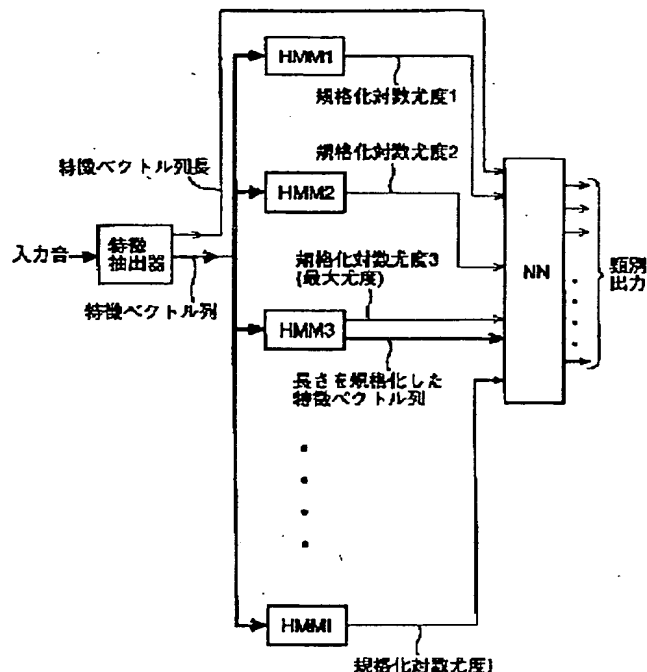
APPLICATION DATE : 27-12-96  
APPLICATION NUMBER : 08350940

APPLICANT : NEC CORP;

INVENTOR : KAJIWARA NOBUKI;

INT.CL. : G10L 3/00 G10L 3/00 G06F 15/18  
G10L 9/10

TITLE : SOUND CLASSIFICATION SYSTEM



ABSTRACT : PROBLEM TO BE SOLVED: To achieve a classification method having a higher classifying ability by inputting to a neural network(NN) the length of a feature vector sequence before standardization and a standardized logarithmic likelihood of each hidden Markov model(HMM) in addition to a standardized feature vector sequence.

SOLUTION: Given input sound is converted into a feature vector sequence by a feature extraction unit. The feature vector sequence is processed according to HMM1-HMM3... to be classified, and standardized logarithmic likelihood of each HMM is determined. The number of feature vectors included in the feature vector sequence, namely, the length of the feature vector sequence, is standardized by using the HMM giving a likelihood having a maximum value (maximum likelihood) among the standardized logarithmic likelihoods. The feature vector sequence with the length standardized, the standardized logarithmic likelihoods determined from each HMM, and the length of the feature vector sequence before standardization are inputted to NN. The NN processes the inputs and outputs an estimated classification result.

COPYRIGHT: (C)1998,JPO

**THIS PAGE BLANK (USPTO)**

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-187183

(43) 公開日 平成10年(1998) 7月14日

(51) Int.Cl. <sup>8</sup>	識別記号	F I
G 1 0 L 3/00	5 3 5	G 1 0 L 3/00 5 3 5
	5 3 1	5 3 1 E
G 0 6 F 15/18	5 2 0	G 0 6 F 15/18 5 2 0 E
G 1 0 L 9/10	3 0 1	G 1 0 L 9/10 3 0 1 C

審査請求 有 請求項の数 2 O L (全 12 頁)

(21) 出願番号 特願平8-350940

(22) 出願日 平成 8 年 (1996) 12 月 27 日

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目 7 番 1 号

(72) 発明者 中谷 正吾

東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

(72) 発明者 梶原 信樹

東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

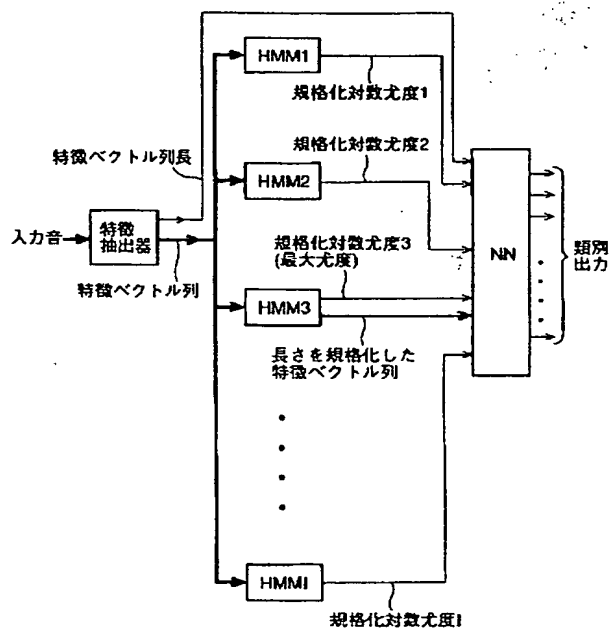
(74) 代理人 弁理士 後藤 洋介 (外 2 名)

(54) 【発明の名称】 音類別方式

(57) 【要約】

【課題】 高い類別能力を有する類別方式を提供することである。

【解決手段】 最大尤度を示すHMMで長さを規格化した特徴ベクトル列と、全カテゴリーのHMMの規格化した対数尤度と、長さを規格化する前の特徴ベクトル列の長さをNNの入力とする。これによって、NNはHMMによる類別結果とNNのみによる類別結果を総合的に判断し類別出力を出す。



**【特許請求の範囲】**

【請求項1】 類別すべきカテゴリに対応する隠れマルコフモデル(HMM)と、類別すべきカテゴリに対応する出力を有する多層パーセプトロンからなり、入力特徴ベクトル列の長さ(ベクトルの数)と、前記隠れマルコフモデルのうち入力特徴ベクトル列に対して最大尤度を示す隠れマルコフモデルを用いて長さを規格化した特徴ベクトル列と、前記各カテゴリに対応する隠れマルコフモデルの入力特徴ベクトル列に対する対数尤度を入力特徴ベクトル列の長さで割ったものを前記多層パーセプトロンの入力とすることを特徴とする音類別方式。

【請求項2】 前記入力特徴ベクトル列の長さ、と、前記した各カテゴリに対応する隠れマルコフモデルの入力特徴ベクトル列に対する対数尤度を入力特徴ベクトル列の長さで割ったものが、多層パーセプトロンの中間層を経ず、多層パーセプトロンの出力層に直接入力することを特徴とする請求項1記載の音類別方式。

**【発明の詳細な説明】****【0001】**

【発明の属する技術分野】本発明は、過渡音(物をたたいたりしたときなどに発生する短い時間内で過渡的に発生する音)や音声を、処理装置を用いて自動的に類別する方式に関するものである。

**【0002】**

【従来の技術】過渡音からその発生源を類別したり、音声から発話された語を認識する方式として、HMM(隠れマルコフモデル)やNN(ニューラルネットワーク)を用いる方式が広く知られている。さらに、HMMやNNを単独で用いた場合よりも高い類別能力を得るため、両者を組み合わせて用いる方法も提案されている。

【0003】図10にHMMとNNを組み合わせた従来の音類別方式の構成を示す。このシステムを用いて類別を行うには、あらかじめシステムに含まれるパラメータを学習によって決めなければならない。まず、以下に学習過程について説明する。

【0004】音データは適当な前処理によって特徴ベクトル列に変換される。類別すべきカテゴリ数をIとする。各カテゴリに属する学習用データを用い、各カテゴリに対応する隠れマルコフモデル、HMM1、HMM2、HMM3、…、HMMIを作成する。次に各々の学習用データを全HMMで処理し、最大の尤度を示すHMMを用いて特徴ベクトル列の伸縮を取り除いた長さを規格化した特徴ベクトル列を作成する。最後にこのようにして得られた長さを規格化した特徴ベクトル列を用いてNNが所望の類別出力を出すように学習させる。

【0005】つぎに類別過程について述べる。まず入力音は特徴ベクトル列に変換され、前記特徴ベクトル列を全HMMに入力する。次に最大の尤度を示すHMMを用いて長さを規格化した特徴ベクトル列を作成する。最後に長さを規格化した特徴ベクトル列をNNに入力し、N

Nの出力で音を類別する。

【0006】図11に前記従来例のNNを示す。これは3層パーセプトロンで、長さを規格化した特徴ベクトル列を入力とし、その音が属するカテゴリに対応する出力のみがアクティブになるように学習される。

【0007】しかしながら、本従来方式ではHMMが特徴ベクトル列の長さの規格化にのみ利用されており、HMMが有する類別能力は有効に使われていない。音データによってはHMMの方がNNより正しい類別結果を示す場合があり、HMMの類別能力の有効利用によって、より高い類別能力を有する方式が望まれる。

**【0008】**

【発明が解決しようとする課題】従来方式の問題点は、HMMが特徴ベクトル列の長さの規格化にのみ用いられており、HMMが有する類別能力はもっぱら規格化に採用するHMMを決めるためだけにしか利用されていないため、HMMが有する類別能力が有効に利用されていないことである。

【0009】音の種類によってはHMMの方がNNより正しい類別結果を示す場合があり、HMMによる類別結果の有効利用が望まれる。

【0010】本発明の目的は、HMMによる類別結果も有効に利用することによって、より高い類別能力を有する音類別方式を提供することにある。

**【0011】**

【課題を解決するための手段】本発明によれば、類別すべきカテゴリに対応する隠れマルコフモデル(HMM)と、類別すべきカテゴリに対応する出力を有する多層パーセプトロンからなり、入力特徴ベクトル列の長さ(ベクトルの数)と、前記隠れマルコフモデルのうち入力特徴ベクトル列に対して最大尤度を示す隠れマルコフモデルを用いて長さを規格化した特徴ベクトル列と、前記各カテゴリに対応する隠れマルコフモデルの入力特徴ベクトル列に対する対数尤度を入力特徴ベクトル列の長さで割ったものを前記多層パーセプトロンの入力とすることを特徴とする音類別方式が得られる。

【0012】さらに、本発明によれば、前記入力特徴ベクトル列の長さ、と、前記した各カテゴリに対応する隠れマルコフモデルの入力特徴ベクトル列に対する対数尤度を入力特徴ベクトル列の長さで割ったものが、多層パーセプトロンの中間層を経ず、多層パーセプトロンの出力層に直接入力することを特徴とする音類別方式が得られる。

**【0013】**

【作用】本発明では、図1に示すように、規格化特徴ベクトル列に加えて、規格化前の特徴ベクトル列の長さおよび各HMMの規格化対数尤度もNNの入力とすることにより、自動的にNNがHMMの類別結果をも考慮して総合的に類別することができる。なお、NNとHMMの類別能力を総合的にどのように考慮して判断するのか

は、NNの学習によって自動的に最適な考慮の仕方を決める。

#### 【0014】

【発明の実施の形態】以下、本発明の第1の実施の形態について図面を参照して詳細に説明する。図1は、本発明の第1の実施の形態の構成を示すブロック図である。まず、任意の与えられた入力音は、特徴抽出器によって特徴ベクトル列に変換される。前記特徴ベクトル列は、類別すべき各カテゴリーに対応する隠れマルコフモデルHMM1、HMM2、HMM3、…、で処理され、各HMMの規格化対数尤度が求められる。前記規格化対数尤度のなかで最大の値を持つ尤度（最大尤度）を与えるHMMを用い、前記特徴ベクトル列に含まれる特徴ベクトルの数すなわち特徴ベクトル列の長さが規格化される。前記長さを規格化された特徴ベクトル列と、前記各HMMから求められた規格化対数尤度と、規格化する前の特徴ベクトル列の長さはNNに入力される。前記NNは前記入力処理し、推定される類別結果を出力する。以下各部について詳細に説明する。

【0015】図1の特徴抽出器は、入力音をその特徴をよく表現するベクトルすなわち特徴ベクトルの系列に変換するものである。本発明は特徴ベクトルとして何を用いてもよいが、特徴ベクトルの典型的なものとしてはパワースペクトルが挙げられる。図2にパワースペクトルを用いた特徴抽出の例を示す。まず、入力音はその特徴を損なわないような間隔でサンプリングされる。サンプリングされた入力音は、時間方向に適当な区間に分けられる。図の例では区間1から区間4までの四区間に分けられている。この区間は互いに重なりあってもよい。次に各区間のサンプリングデータに高速フーリエ変換（FFT）を施し、その結果得られた離散的フーリエ係数の絶対値の自乗をとって各区間に対応する離散的パワースペクトルを求める。前記離散的パワースペクトルを成分とするベクトルが区間 $i$ に対応する特徴ベクトル $x_i$ （ $i=1\sim 4$ ）であり、元の入力音はこのようにして得られた特徴ベクトルの列 $X=x_1, x_2, x_3, x_4$ に変換される。

【0016】HMMは状態間の確率的遷移と、各状態における特徴ベクトルの確率的出力を特徴とするモデルである。一つのHMMは、最初に状態 $i$ に居る確率 $\{\pi_i \mid i=1\sim N\}$ と、状態 $i$ から状態 $j$ への遷移確率 $\{a_{ij} \mid i, j=1\sim N\}$ と、状態 $i$ で特徴ベクトル $x$ を出力する確率密度 $\{b_i(x) \mid i=1\sim N\}$ で指定される。ここで、 $N$ はHMMの状態数である。

【0017】与えられた特徴ベクトル列 $X=x_1, x_2, x_3, \dots, x_T$ をHMMが出力する確率すなわち尤度は図3に示すViterbiアルゴリズムで求められる（ここで求められるのは厳密には確率そのものではなく確率に比例する量であるが、簡単のため確率と呼ぶことにする。）。図3において、ステップV1は最初に状態 $i$ に

居る特徴ベクトル $x_1$ を出力する確率 $\xi_1(i)$ の計算と、最適状態遷移経路を保持するための変数の初期化（第2式）を行っている。このステップV1はすべての状態（ $i=1\sim N$ ）に渡って行われる。ステップV2は、特徴ベクトル列 $x_1, x_2, x_3, \dots, x_t$ を出力して状態に至る状態遷移経路のうち最大の特徴ベクトル列出力確率を与えるものを求めるもので、 $\xi_t(i)$ はその出力確率、 $\psi_t(i)$ はその経路において $x_{t-1}$ を出力したときの状態を表す。以下に示す数1は $i$ を1から $N$ まで変えたときの[ ]内の式の最大値である。

#### 【0018】

【数1】

$$\max_{1 \leq i \leq N} [ ]$$

以下に示す数2はその最大値を与える $i$ の値である。

#### 【0019】

【数2】

$$\arg \max_{1 \leq i \leq N} [ ]$$

ステップV2はすべての状態（ $j=1\sim N$ ）に渡って行われ、また第2番目の特徴ベクトルから最後（第 $T$ 番目）の特徴ベクトルにわたって（ $t=2\sim T$ ）逐次実行される。ステップV3は特徴ベクトル列 $X=x_1, x_2, x_3, \dots, x_T$ を出力する状態遷移経路のうち出力確率が最大である経路すなわち最適状態遷移経路の出力確率 $P$ を求めるものである。 $q_T$ は最適状態遷移経路の最終状態である。ステップV4は、 $q_T$ からはじめて、最適状態遷移経路を逆方向に辿るもので、これによって最適状態遷移経路 $q_1, q_2, q_3, \dots, q_T$ が求められる。

【0020】HMMのみでも次のように類別能力がある。まず、類別すべき各カテゴリーに対応するHMMをあらかじめ学習によって決めておく。次に、任意に与えられた特徴ベクトル列を各HMMが出力する確率（尤度）を前記Viterbiアルゴリズムで求める。前記尤度のうち最大の尤度を与えるHMMに対応するカテゴリーに前記特徴ベクトル列が属すると判断することにより類別がなされる。

【0021】次に、HMMを決めるパラメータ $\{\pi_i \mid i=1\sim N\}$ 、 $\{a_{ij} \mid i, j=1\sim N\}$ 、 $\{b_i(x) \mid i=1\sim N\}$ を決める方法である学習について説明する。

【0022】図4に学習の初期過程の概要図を示す。まず、「ドアのノック音」、「靴音」等類別したい各カテゴリーに対応する学習用音データを十分な数用意する（例えば、同じノック音でも、叩く強さ、ドアの質、気温、湿気等により様々になるが、それら種々のデータをノック音として一つのカテゴリーの学習用データとする。）。次に、各音データを前述の特徴抽出器で処理して学習用特徴ベクトル列に変換する。これらの学習用特徴ベクトル列を図5に示すK-meansアルゴリズム

で処理し各カテゴリーに対応するHMMのパラメータを決める。以下にK-Meansアルゴリズムについて説明する。

【0023】まず、各カテゴリーに属する特徴ベクトルをいくつかのクラスターに分類する(図4のクラスタリング)。ここで、各カテゴリーのクラスター数はそのカテゴリーにおける最長学習用特徴ベクトル列の長さと同じにする(これは一般に各カテゴリーごとに異なる)。ただし特徴ベクトル列がT個の特徴ベクトルから成る場合、Tをその特徴ベクトル列の長さと呼ぶ。特徴ベクトルのクラスタリングは種々の方法が適用可能であるが、以下に典型的なクラスタリング法を記す。

【0024】まず、乱数を用いて各成分を生成したN個のランダムな特徴ベクトルを作る。ここでNは最長学習用特徴ベクトル列の長さである。これらのベクトルはセントロイドと呼ばれ、各セントロイドに一つのクラスターが対応付けられる。次に、カテゴリー内の特徴ベクトルと各セントロイドとの距離を計算し、もっとも距離が短いセントロイドに対応するクラスターにその特徴ベクトルが属するものとする。このようにして、カテゴリー内の全ての特徴ベクトルをクラスターに分類することができる。ここでベクトル間の距離はユークリッド距離を用いる。各クラスターには1からNまで順に番号を付する。K-meansアルゴリズムではクラスターはHMMの状態に一对一に対応しており、したがってこの番号はHMMの状態を表すインデックスとみなすことができる。このようにクラスタリングの後、各特徴ベクトルに一意に状態が対応付けられ、したがって各特徴ベクトル列には一つの状態遷移経路が対応付けられる。

【0025】上述した学習の初期過程におけるクラスタリングが、K-meansアルゴリズム図5のステップ

$$b_i(x) = \frac{1}{\sqrt{(2\pi)^C \det(V_i)}} \exp\left(-\frac{1}{2}(x-\mu_i)^T V_i^{-1}(x-\mu_i)\right)$$

ここで、Cは特徴ベクトルの次元、 $\det(V_i)$ は $V_i$ の行列式、 $V_i^{-1}$ は $V_i$ の逆行列である。

【0027】ステップK6においては、前述のViterbiアルゴリズムを用いて各学習用特徴ベクトル列の最適状態遷移経路を求める。この最適状態遷移経路がもしステップK1で与えられた状態遷移経路と異なる場合、最適状態遷移経路にしたがって特徴ベクトルの再クラスタリングを行う。すなわち、任意のtに対して、第t番目の特徴ベクトル $x_t$ が最適状態遷移経路の第t番目の状態 $q_t$ に対応するクラスターに分類されるように、各特徴ベクトル列の特徴ベクトルをクラスタリングし直す。もし、ステップK6で再クラスタリングが必要なければ、すなわちすべての学習用特徴ベクトル列についてステップK1で与えられた状態遷移経路とViterbiアルゴリズムで求めた最適状態遷移経路が一致していればK-meansアルゴリズムは終了する。もし

K1である。K-meansアルゴリズムは各カテゴリーごとに独立に施される。図5及び以下の説明は一つのカテゴリーに関するものとする。図5ではクラスター数すなわち状態数をNとしている。ステップK2、K3の総和は、クラスターiに属する特徴ベクトル $x_t$ について加算することを意味する。 $N_i$ はクラスターiに属する特徴ベクトルの数である。ステップK2の $\mu_i$ はクラスターiの平均ベクトルである。ステップK3の $\Gamma(\cdot)$ は $(\cdot)$ 内のベクトルを転置するすなわち行ベクトルを列ベクトルに変換することを意味する。ただし、特徴ベクトル $x_t$ は行ベクトルである。ステップK3の $V_i$ はクラスターiの分散行列である。ステップK4はHMMの初期状態がiである確率 $\pi_i$ を与えるものである。ステップK4の右辺の分子はカテゴリー内の学習用特徴ベクトル列(例えば $X_1, X_2, X_3, \dots$ )のうち、第一番目の特徴ベクトル( $x_1$ )がクラスターiに属するものの数である。ステップK5は、状態iから状態jへの遷移確率 $a_{ij}$ を与えるものである。右辺分子は、一つの特徴ベクトル列(例えば $X_1, X_2, X_3, \dots$ )のなかでクラスターiに含まれる特徴ベクトル $x_t$ があつて、次にクラスターjに含まれる特徴ベクトル $x_{t+1}$ が来るような並びの数(tは任意)を数え上げ、そのようにして得られた数をすべての特徴ベクトル列について足し上げたものである。このようにしてすべての状態(i, j = 1~N)に関するパラメータ $\mu_i, V_i, \pi_i, a_{ij}$ を与えることにより、一つのHMMが指定される。なお、状態iで特徴ベクトルxを出力する確率密度はガウシアン分布を想定し、以下の数3に示す数式で計算される。

【0026】

【数3】

ステップK6において再クラスタリングがなされた場合は、図5のステップK1から後のステップを再度実行する。次にステップK6に達したとき、そこであらためて求めた最適状態遷移経路と前回求めた最適状態遷移経路とが異なれば、新しく求めた最適状態遷移経路にしたがって再クラスタリングを行い、再び図5のステップK1の後から実行する。もし、再クラスタリングがなければK-meansアルゴリズムは終了する。このように、ステップK6で再クラスタリングがなくなるまでステップK1より後の処理を繰り返す。K-meansアルゴリズムが終了した時点で、HMMを定めるパラメータ $\mu_i, V_i, \pi_i, a_{ij}$ (i, j = 1~N)が決まり、そのカテゴリーに対するHMMが求まったことになる。このようにして、すべてのカテゴリーに対応するHMMを求めるとHMMの学習過程は終了する。

【0028】次に、HMMによる特徴ベクトル列の長さ

の規格化について説明する。過度音や音声は、その時々で音の長さが様々に変化する。NN(ニューラルネット)はそのような音長の変化に対して弱いため、HMMを使って音長の変化を取り除いた長さを規格化した特徴ベクトル列を作り、それをNNの入力とする。この長さの規格化は次のように行う。まず、与えられた特徴ベクトル列の最適状態遷移経路をViterbiアルゴリズムで求める。次に、その最適状態遷移経路のなかで同じ状態が連続して続く並びがあった場合、その並びに対応する特徴ベクトル列をそれらのベクトルの平均ベクトルで置き換える。

【0029】また、HMMは、本来の音全体の土地中から音が始まったり、途中で音が切れたりした場合でもある程度類別する能力がある。このような欠落のある特徴ベクトル列の場合、その最適状態遷移経路と該当カテゴリ内の他の特徴ベクトル列の最適状態遷移経路とを比較して欠けている状態部分に、背景音に対応する特徴ベクトルあるいはすべての成分が0である特徴ベクトルを補う。このようにしてカテゴリ内の総ての特徴ベクトル列の長さを揃える。さらに、異なるカテゴリ間の特徴ベクトル列の長さも揃えるため、全カテゴリの規格化特徴ベクトル列を最も長い規格化特徴ベクトル列の長さにあわせる。ここで、長さの不足分には背景音に対応する特徴ベクトルあるいはすべての成分が0である特徴ベクトルを、特徴ベクトル列の先頭に(あるいは最後にでもよい)補う。このようにして、最も長い学習用規格化特徴ベクトル列の長さ(以下これを $T_L$ とする)にすべての特徴ベクトル列の長さをそろえる。このようにして出来た特徴ベクトル列を以下規格化特徴ベクトル列と呼ぶ。

【0030】次にNNについて述べる。本発明の第1の実施の形態では、図6に示す3層パーセプトロンを用いる。図6の黒丸●は入力をそのまま出力し次段の各入力へ分配するものである。図6の大きい白丸○は、図7に示すニューロンを表す。図7は、ニューロンjが、入力 $s_k$  ( $k=1\sim K$ )にそれぞれ $w_{ik}$ の重みを掛け、それらを加算したものと閾値 $\theta_j$ との差をとったものを非線形関数 $f(\cdot)$ の引数とし、そのときの非線形関数 $f(\cdot)$ の値を出力 $S_j$ とするものであることを示す。この非線形関数 $f(\cdot)$ は以下の数4に示されるように通常はシグモイド関数が用いられる。

【0031】

【数4】

$$f(x) = \frac{1}{1 + e^{-x}}$$

上記した第1の実施の形態では、規格化特徴ベクトル列、全HMMの規格化対数尤度、および規格化前の特徴ベクトル列の長さをNNの入力とする。規格化対数尤度とは、図3のViterbiアルゴリズムで求めた尤度Pの対数を取り、それを規格化前の特徴ベクトル列の長

さで割ったもののことである。対数をとることによりNNの入力のダイナミックレンジを抑えることができ、また規格化前の特徴ベクトル列の長さで割ることにより、異なる長さをもつ特徴ベクトル列間での対等な尤度比較が可能となる。

【0032】特徴ベクトルの次元をCとすると、規格化特徴ベクトル列は全部で $CT_L$ の成分からなる。また、類別カテゴリの数をIとすると、NNの全入力数Kは $CT_L + I + 1$ である。NNの出力層は類別カテゴリの数I個のニューロンからなり、特徴ベクトル列が属するカテゴリに対応する出力のみがアクティブ(1)になり、他の出力はインアクティブ(0)になるように学習される。中間層ニューロンの数Jは、少ない数から類別結果を見ながら徐々に増やしていき、所望の類別能力が得られるところで止めることで決定する。

【0033】次にNNの学習アルゴリズムであるバックプロパゲーションアルゴリズム(図8)について説明する。図8で、mは学習サンプルを表すインデックス(カテゴリの枠を取り払って全学習用特徴ベクトル列にわたって付けた通し番号)、kは入力層のニューロン(入力をそのまま出力する)を表すインデックス、jは中間層のニューロンを表すインデックス、iは出力層のニューロンを表すインデックス、 $s_k^m$ は学習サンプルmのk番目の成分、 $w_{ij}$ と $\theta_j$ はそれぞれ中間層ニューロンjの重みと閾値、 $S_j^m$ は学習サンプルmに対応する中間層ニューロンjの出力、 $W_{ij}$ と $\Theta_i$ はそれぞれ出力層ニューロンiの重みと閾値、 $L_i^m$ は学習サンプルmに対する教師出力のi番目の成分で、その学習サンプルが属するカテゴリに対応する成分のみが1で他の成分は0である。Iは出力層ニューロンの数(したがって類別カテゴリの数)、Jは中間層ニューロンの数、Kは入力層ニューロンの数 $CT_L + I + 1$ である。ニューロンのパラメータ $w_{ik}$ 、 $\theta_j$ 、 $W_{ij}$ 、 $\Theta_i$ は、乱数によって適当な初期値を与える。図8で、ステップB1及びB2は学習サンプルmに対する中間層ニューロンjの出力を計算しており、これをすべての中間層ニューロンについて求める。ステップB3は中間層ニューロンの出力から出力層ニューロンiにおける非線形関数 $f(\cdot)$ の引数を計算する。ステップB4は、出力層ニューロンのパラメータ $W_{ij}$ および $\Theta_i$ の改良値を計算するのに必要なパラメータ $D_i^m$ を求めるものである。このパラメータはすべての出力層ニューロンについて計算する。ここで $f(\cdot)$ は $f(\cdot)$ の微分である。ステップB5は、中間層ニューロンのパラメータ $w_{ik}$ および $\theta_j$ の改良を計算するのに必要なパラメータ $d_j^m$ を求めるもので、これをすべての中間層ニューロンについて計算する。以上の処理をすべての学習サンプルについて行う。ステップB6は中間層ニューロンの重みパラメータをステップB7は中間層ニューロンの閾値パラメータをそれぞれ改良するもので、右辺の $w_{ik}$ 、 $\theta_j$ は従来値、左辺のそれは改良値

である。この改良をすべての中間層のパラメータについて行う。 $\epsilon$ は正の定数である。ステップB8は出力層ニューロンの重みパラメータをステップB10は出力層ニューロンの閾値パラメータをそれぞれ改良するもので、右辺の $w_{ij}$ 、 $\theta_i$ は従来値、左辺のそれは改良値である。この改良をすべての出力層のパラメータについて行う。中間層および出力層のすべてのパラメータの従来値と改良値の差が十分小さくなれば収束したものとして学習を終了し、そうでなければ改良されたパラメータを使って本アルゴリズムを最初から再実行する。以上のバックプロパゲーションアルゴリズムによってNNのパラメータが決定される。上記のように学習によってHMMとNNのパラメータが決定されたのち、図1に示す処理によって類別を行うことができる。類別は次の手順で行う。

【0034】まず、任意の与えられた入力音は、特徴抽出器によって特徴ベクトル列に変換される。次に前記特徴ベクトル列は、各カテゴリーに対応する隠れマルコフモデルHMM1、HMM2、HMM3、...、で処理され規格化対数尤度が求められる。前記規格化対数尤度のなかで最大の値を与えるHMMを用い、前記特徴ベクトル列の長さが規格化される。前記規格化特徴ベクトル列と、前記各HMMから求められた規格化対数尤度と、長さを規格化する前の特徴ベクトル列の長さをNNに入力する。前記NNは前記入力音を処理し、特徴ベクトル列が属すると推定されるカテゴリーに対応するニューロンの出力をアクティブにする。これが、本システムが判断する類別結果となる。

【0035】次に、本発明の第1の実施の形態の効果について説明する。一般に類別結果は、特徴ベクトル列の種類や長さに依存してNNの類別結果が正しかったりHMMの類別結果が正しかったりする。本発明の第1の実施の形態では、規格化特徴ベクトル列のみならず、規格化前の特徴ベクトル列の長さおよび各HMMの規格化対数尤度もNNの入力とすることにより、NNがHMMの類別結果をも考慮して総合的に判断するようになっている。NNとHMMの類別能力を総合的にどのように考慮して最終的な類別を行うかも、NNの学習によって自動的に最適な考慮の仕方が決められる。これによって、従来のようにNNのみあるいはHMMのみで類別した場合にくらべて、より高い類別能力を提供することができる。

【0036】次に、本発明の第2の実施の形態について説明する。本発明の第2の実施の形態では、図1におけるNNとして図9に示したものをを用いる。すなわちHMMの類別結果である各HMMの規格化対数尤度、およびHMMとNNの類別結果の考慮の仕方に重要な影響をもつパラメータである規格化前の特徴ベクトル列の長さを、直接出力層に入力する。

【0037】学習アルゴリズムは基本的に第1の実施の

形態におけるアルゴリズム図8と同じであるが、各HMMの規格化対数尤度および規格化前の特徴ベクトル列の長さの入力は、入力層ニューロンには含めず中間層ニューロンに含める。ただし、各HMMの規格化対数尤度および規格化前の特徴ベクトル列の長さの入力に対応するニューロンは入力層ニューロンとおなじく入力をそのまま出力するものである。

【0038】したがって本発明の第2の実施の形態では図8の入力層ニューロンの数Kは規格化特徴ベクトル列の全成分数 $CT_L$ とする。また、ステップB3およびB9におけるJは $J+I+1$ で置き換え、 $j=J+1\sim J+I+1$ における $S_j$ は各HMMの規格化対数尤度および規格化前の特徴ベクトル列の長さであるとする。

【0039】本発明の第2の実施の形態は、第1の実施の形態の効果に加えて、学習時間が第1の実施の形態より短くなるという効果を有する。これは、規格化前の特徴ベクトル列の長さおよび各HMMの規格化対数尤度が、中間層ニューロンを bypass 直接出力層に繋がっているため、第1の実施の形態にくらべてより少ない中間層ニューロン数で済むためである。

【0040】

【発明の効果】第1の効果は、従来方式に比べてより高い類別能力を提供できることである。

【0041】その理由は、規格化特徴ベクトル列に加えて、HMMによる類別結果（各HMMの規格化対数尤度）および規格化前の特徴ベクトル列の長さもNNの入力として、HMMとNNの類別能力を総合的に考慮して類別を行うためである。

【0042】第2の効果は、NNとHMMの類別能力を総合的に考慮して最適な類別する仕方が、自動的に得られることである。

【0043】その理由は、NNの学習によって両者の最適な考慮の仕方が自動的に獲得されるからである。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態を示すブロック図である。

【図2】特徴抽出の概要図である。

【図3】Viterbiアルゴリズムの流れ図である。

【図4】HMMの学習の初期過程の概念図である。

【図5】K-meansアルゴリズムの流れ図である。

【図6】本発明の第1の実施の形態のニューラルネットを示した図である。

【図7】ニューロンの動作の概念図である。

【図8】バックプロパゲーションアルゴリズムの流れ図である。

【図9】本発明の第2の実施の形態のニューラルネットを示した図である。

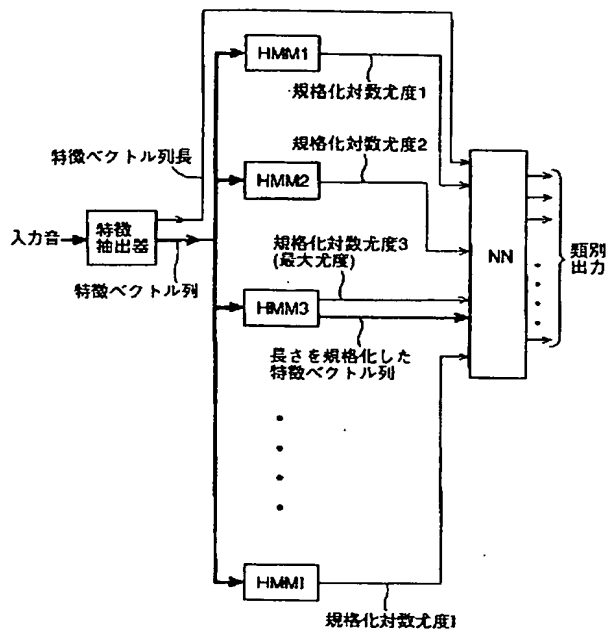
【図10】従来技術の一実施の形態を示したブロック図である。

【図11】従来技術のニューラルネットを示した図であ

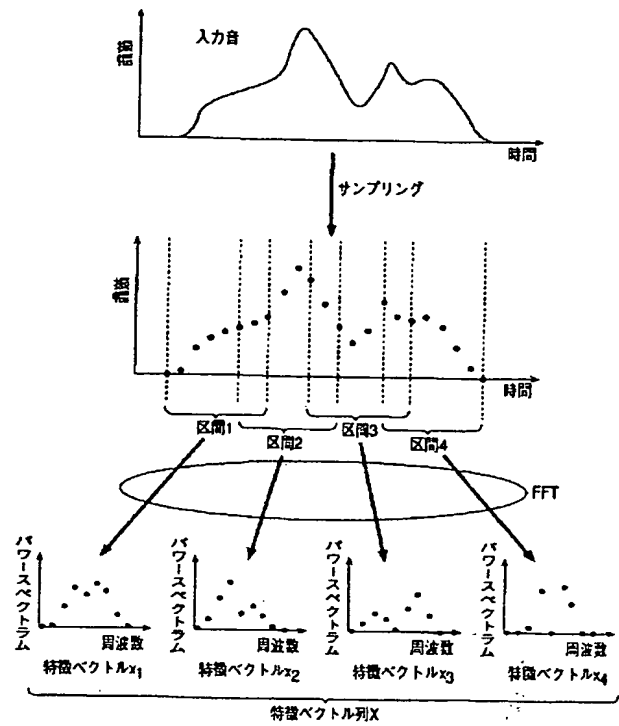


る。

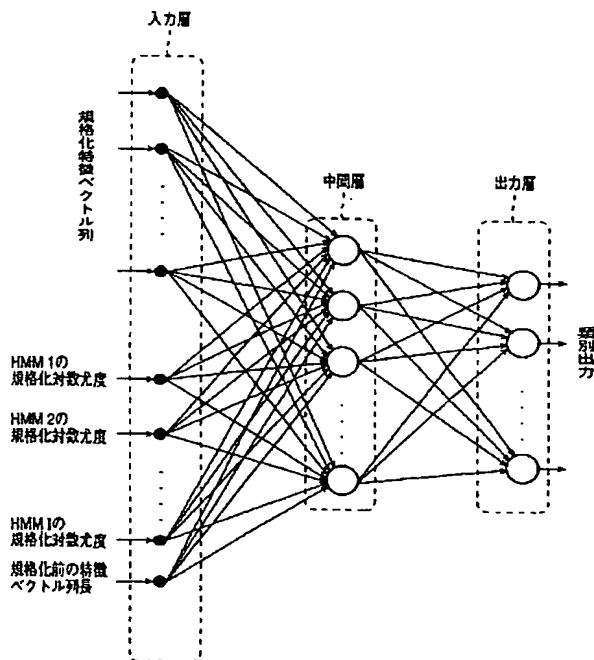
【図1】



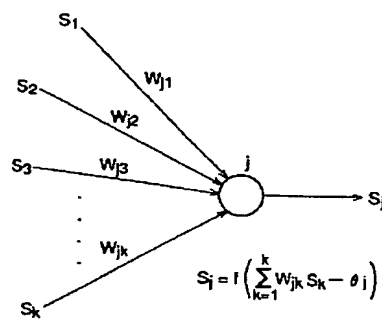
【図2】



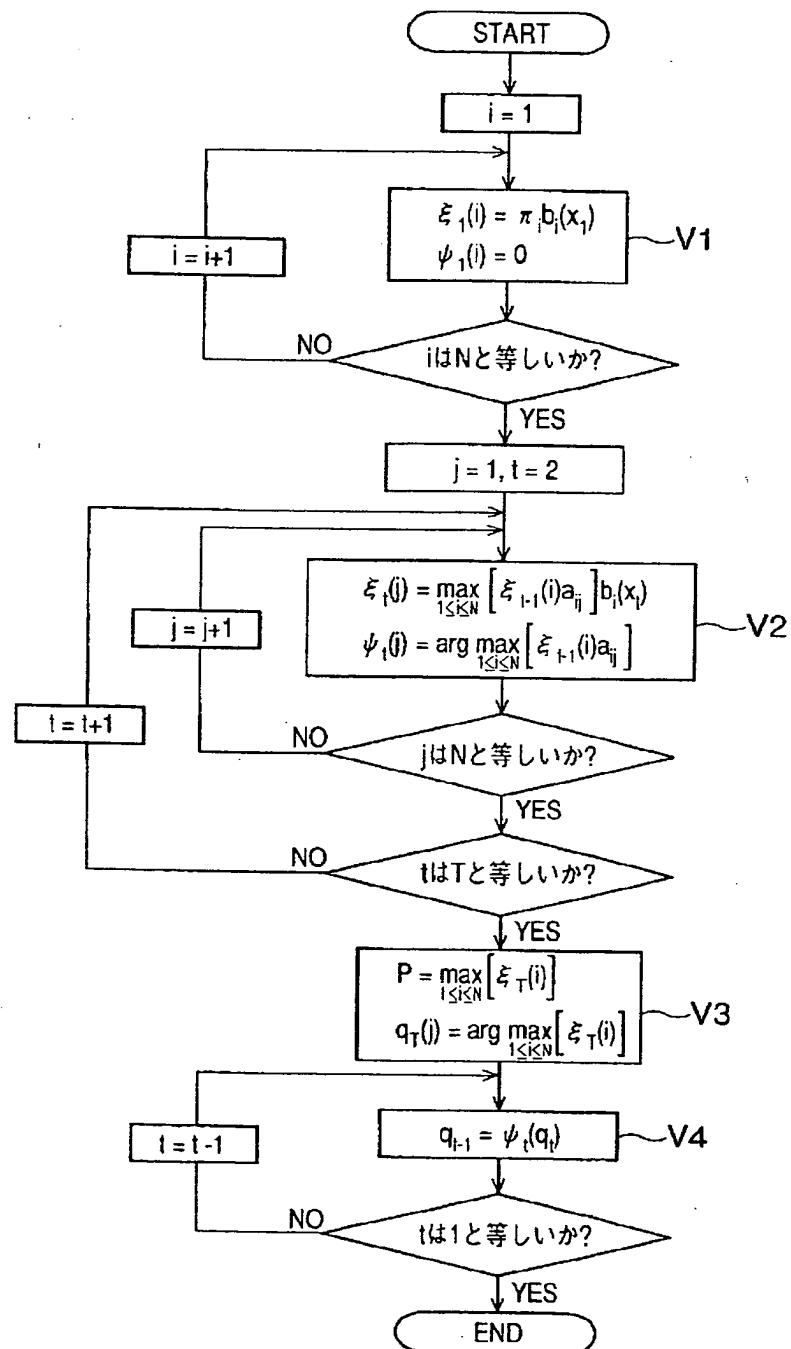
【図6】



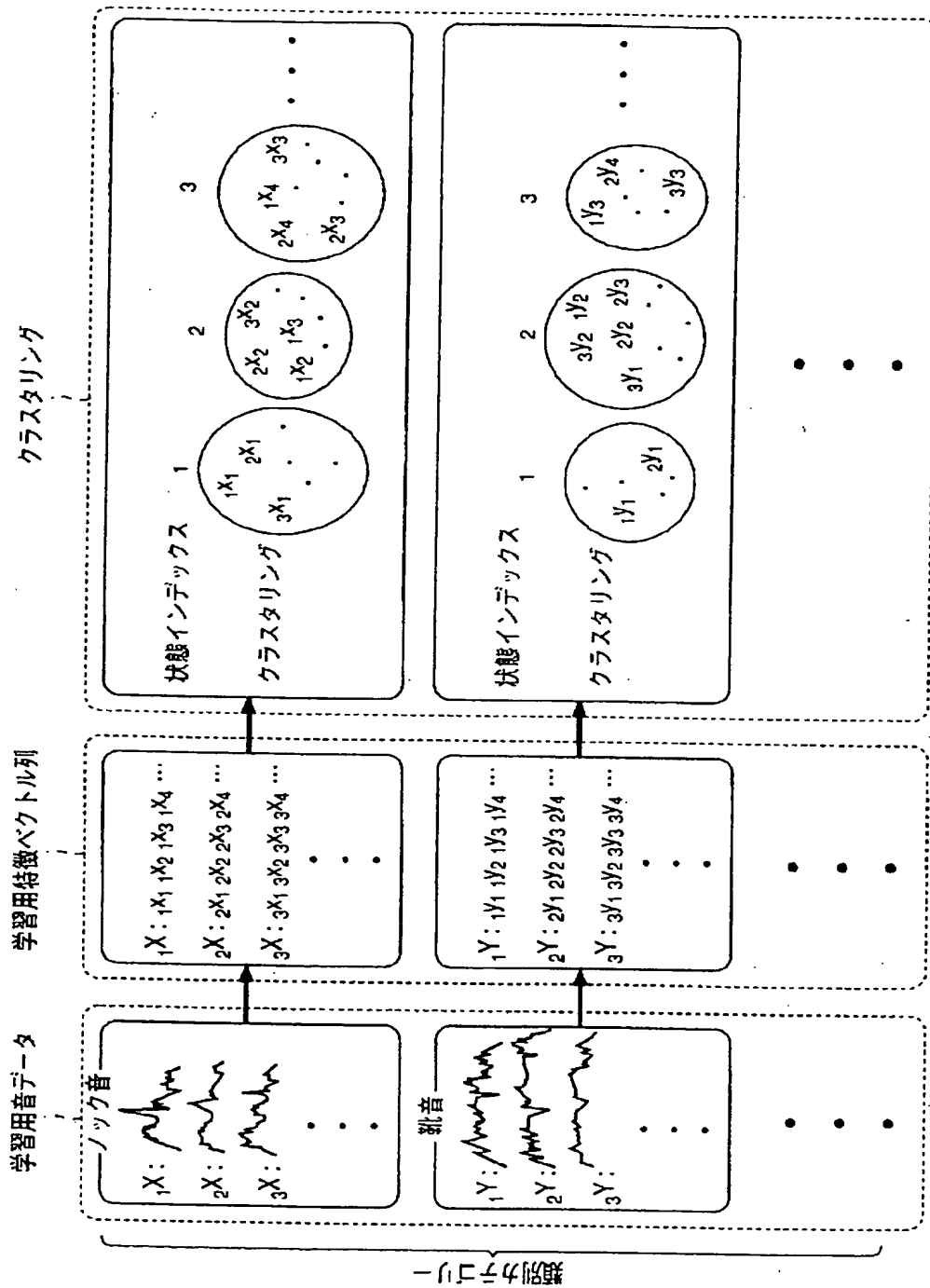
【図7】



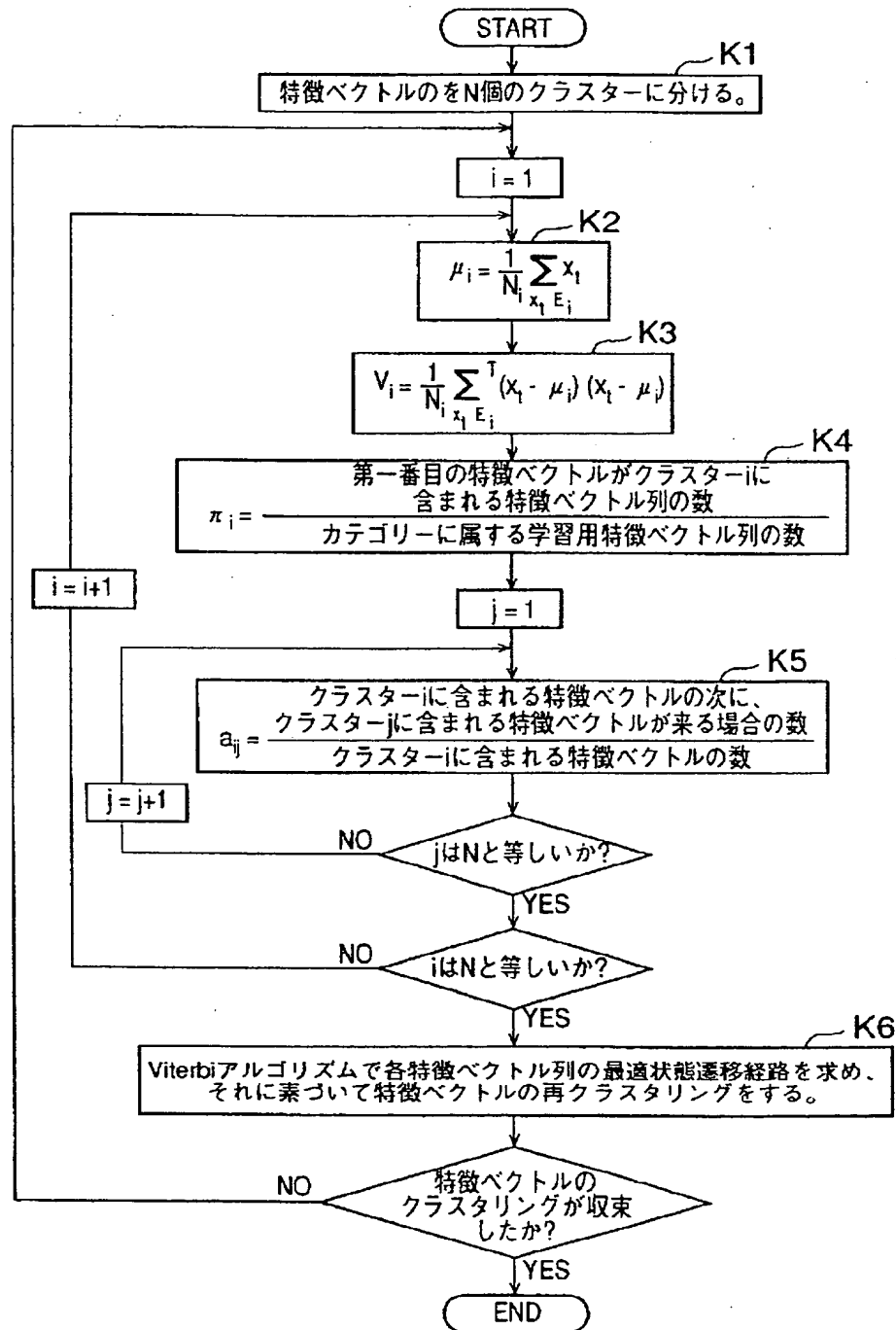
【図3】



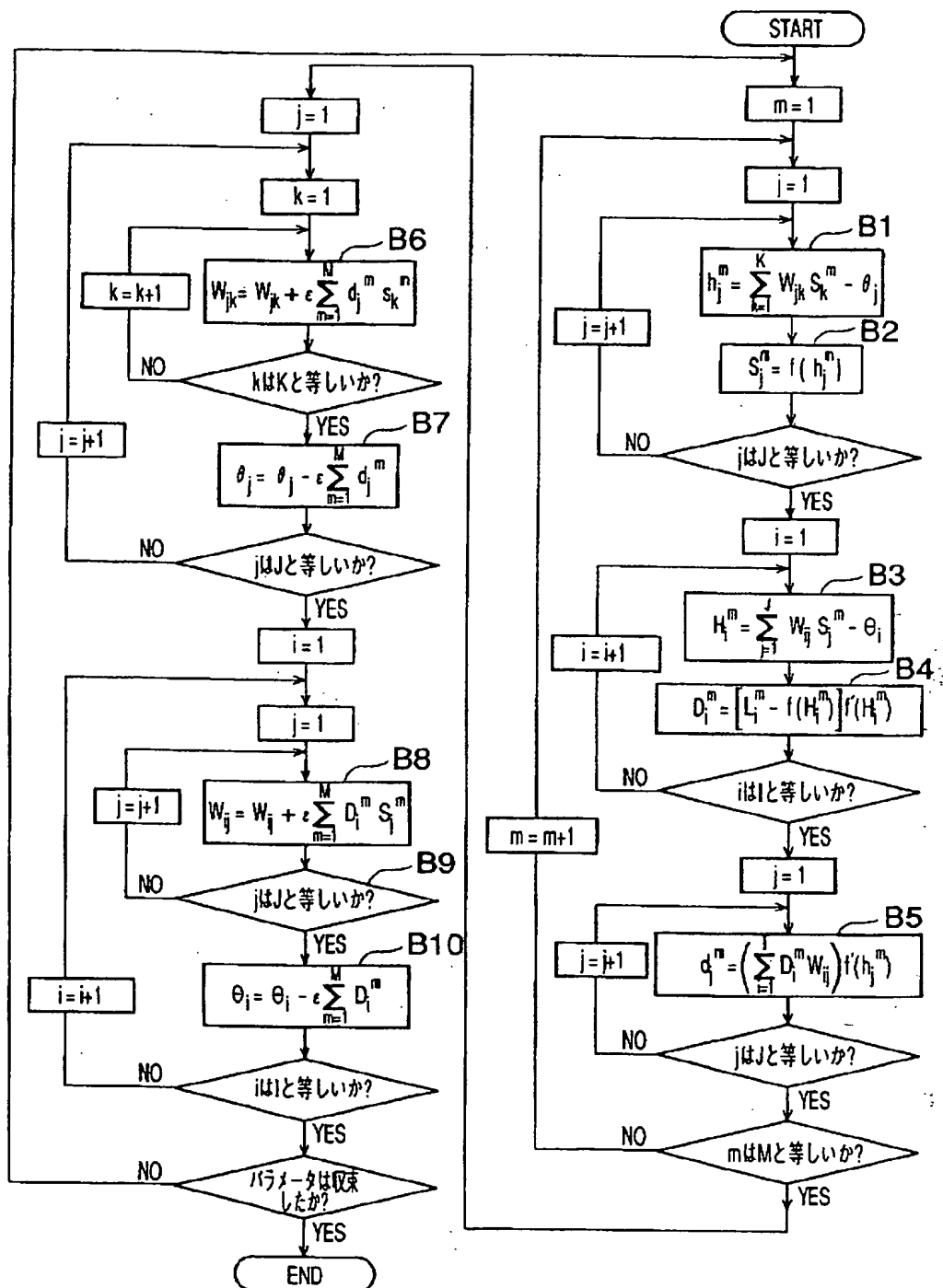
【図4】



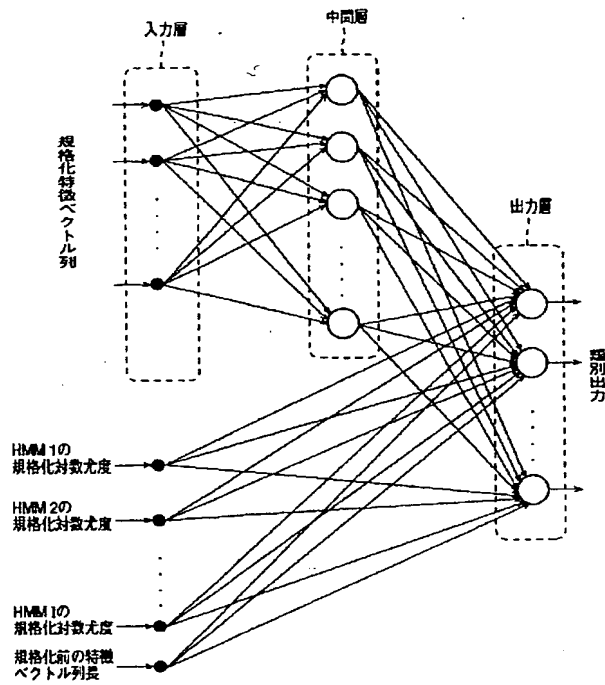
【図5】



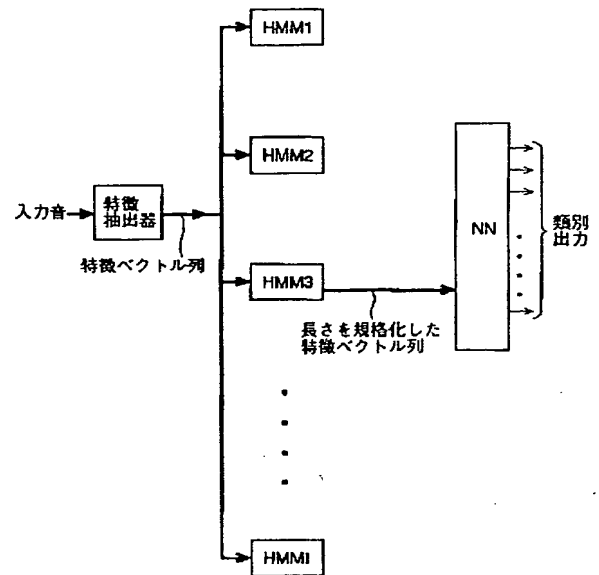
【図8】



【図9】



【図10】



【図11】

